CONVERGENCE OF SAMPLE PATHS OF NORMALIZED SUMS OF INDUCED ORDER STATISTICS

by

P. K. Bhattacharya
University of Arizona and University of Minnesota

Technical Report No. 169

January 1972

University of Minnesota
Minneapolis, Minnesota

CONVERGENCE OF SAMPLE PATHS OF NORMALIZED SUMS

OF INDUCED ORDER STATISTICS[1]

1. Introduction.

$(X_1, Y_1), (X_2, Y_2), \ldots$ are independent two-dimensional random vectors each distributed as (X, Y). Let X_{nk} be the k^{th} order statistic obtained from $X_1, \ldots, X_n$. If the marginal distribution of X is continuous, $X_{n1} < \ldots < X_{nn}$ with probability 1 and we can unambiguously define induced order statistics $Y_{n1}, \ldots, Y_{nn}$ as $Y_{nk} = Y_j$ if $X_{nk} = X_j$. Let $m(x)$ denote the conditional expectation and $\sigma^2(x)$ the conditional variance of Y given $X = x$, and let $\psi(t) = \int_{\infty}^{F^{-1}(t)} \sigma^2(x)dF(x)$, $0 \leq t \leq 1$. The main result in this paper concerns the limiting behavior of the sample paths of

$$\{S_{nk} = \sum_{j=1}^{k} (Y_{nj} - m(X_{nj})),\ k = 1, \ldots, n\}.$$

By means of a Skorokhod-type embedding (see Skorokhod (1961), p. 163) of $\{S_{nk}, k = 1, \ldots, n\}$ on Brownian Motion paths (Theorem 1), it is shown that under certain conditions there are processes $\{\xi^{(n)}(t), 0 \leq t \leq 1\}$ for each n and a Brownian Motion $\{\xi(t), t \geq 0\}$ on a common probability space so that $\{\xi^{(n)}(t), 0 \leq t \leq 1\}$ has the same distribution as $\{S_{n,[nt]}/\sqrt{n\psi(1)}, 0 \leq t \leq 1\}$ and $\sup_{0\leq t\leq 1} |\xi^{(n_j)}(t) - \xi(\psi(t)/\psi(1))| \to 0$ a.s. for sufficiently rapidly increasing subsequences $\{n_j\}$ (Theorem 2). This yields an invariance principle similar to Donsker's (1951). In particular, the asymptotic distribution of $\sup_{0\leq t\leq 1} |S_{n,[nt]}|/\sqrt{n\psi(1)}$ is the same as the distribution of $\sup_{0\leq t\leq 1} |\xi(t)|$. Large sample tests for a specified regression function are obtained from these results.

Preliminaries.

Let F denote the marginal cdf of X and G_x the conditional cdf of Y given $X = x$. We assume that F and $\{G_x\}$ satisfy the following conditions.

[1]AMS 1970 subject classifications: Primary 60F99, secondary 60E?0. Key words and phrases: induced order statistics, Skorokhod embedding, invariance principle, test for regression function.

Condition 1.

F is continuous.

Condition 2.

$\beta(x) = E[\{Y - m(x)\}^4 | X = x]$ is bounded above by some constant B on $(-\infty, \infty)$.

Condition 3.

$\sigma^2(x) = E[\{Y - m(x)\}^2 | X = x]$ is of bounded variation on $(-\infty, \infty)$.

We define function $\psi(t)$ and $\psi_n(t)$ on $[0, 1]$ as follows. Let $F^{-1}(t) = \inf\{x | F(x) = t\}$, $0 < t < 1$, $F^{-1}(0) = -\infty$ and $F^{-1}(1) = +\infty$. Then

$$(1) \qquad \psi(t) = \int_{\infty}^{F^{-1}(t)} \sigma^2(x) dF(x)$$

and

$$(2) \qquad \psi_n(t) = \begin{cases} n^{-1} \sum_{k=1}^{[nt]} \sigma^2(X_{nk}) = \int_{-\infty}^{X_{n,[nt]}} \sigma^2(x) dF_n(x), & 1/n \leq t \leq 1 \\ 0, \quad 0 \leq t < 1/n , \end{cases}$$

where F_n is the empirical cdf of $X_1, \ldots, X_n$ and $[a]$ is the largest integer $\leq a$.

We conclude this section with two lemmas. Lemma 1 gives the conditional distribution of $Y_{n1}, \ldots, Y_{nn}$ given $X_1, \ldots, X_n$ and Lemma 2 establishes the almost sure uniform convergence of $\psi_n(t)$ to $\psi(t)$.

Lemma 1.

Under condition 1, for every n and almost all $(X_1, \ldots, X_n)$, $Y_{n1}, \ldots, Y_{nn}$ are conditionally independent given $X_1, \ldots, X_n$ with conditional cdf's $G_{X_{n1}}, \ldots, G_{X_{nn}}$ respectively.

Proof:

For any $x_1, \ldots, x_n$ no two of which are equal, let $\lambda(n,k; x_1, \ldots, x_n) = j$ if x_j is the k^{th} smallest among $x_1, \ldots, x_n$. Obviously, $\lambda(n,k; x_1, \ldots, x_n)$,

$k = 1, \ldots, n$ is a permutation of $1, \ldots, n$ and $x_{\lambda(n,1;x_1,\ldots,x_n)} < \ldots < x_{\lambda(n,n;x_1,\ldots,x_n)}$. Since by condition 1, $X_1, \ldots, X_n$ are all distinct with probability one, $\lambda(n,k; X_1, \ldots, X_n)$, $k = 1, \ldots, n$ are defined for almost all $(X_1, \ldots, X_n)$ and $X_{nk} = X_{\lambda(n,k;X_1,\ldots,X_n)}$, $Y_{nk} = Y_{\lambda(n,k;X_1,\ldots,X_n)}$. To prove the lemma, it suffices to show that

$$(3) \qquad P[Y_{nk} \leq y | X_1 = x_1, \ldots, X_n = x_n] = G_{x_{\lambda(n,k;x_1,\ldots,x_n)}}(y)$$

and

$$(4) \qquad P[Y_{n1} \leq y_1, \ldots, Y_{nn} \leq y_n \mid X_1 = x_1, \ldots, X_n = x_n] = \prod_{k=1}^{n} G_{x_{\lambda(n,k;x_1,\ldots,x_n)}}(y_k$$

for arbitrary $x_1, \ldots, x_n$ no two of which are equal. Now for such $x_1, \ldots, x_n$,

$$P[Y_{nk} \leq y \mid X_1 = x_1, \ldots, X_n = x_n]$$

$$= P[Y_{\lambda(n,k;X_1,\ldots,X_n)} \leq y \mid X_1 = x_1, \ldots, X_n = x_n]$$

$$= P[Y_{\lambda(n,k;x_1,\ldots,x_n)} \leq y \mid X_1 = x_1, \ldots, X_n = x_n]$$

$$= P[Y_{\lambda(n,k;x_1,\ldots,x_n)} \leq y \mid X_{\lambda(n,j;x_1,\ldots,x_n)} = x_{\lambda(n,j;x_1,\ldots,x_n)},\ j=1, \ldots, n].$$

Since $Y_{\lambda(n,k;x_1,\ldots,x_n)}$ is independent of $\{X_{\lambda(n,j;x_1,\ldots,x_n)},\ j \neq k\}$, (3) is proved. To prove (4) we use the same arguments to get

$$P[Y_{n1} \leq y_1, \ldots, Y_{nn} \leq y_n \mid X_1 = x_1, \ldots, X_n = x_n]$$

$$= P[Y_{\lambda(n,k;x_1,\ldots,x_n)} \leq y_k,\ k = 1, \ldots, n \mid X_{\lambda(n,k;x_1,\ldots,x_n)} = x_{\lambda(n,k;x_1,\ldots,x_n)},\ k=1,\ldots,n]$$

$$= \prod_{k=1}^{n} P[Y_{\lambda(n,k;x_1,\ldots,x_n)} \leq y_k \mid X_{\lambda(n,k;x_1,\ldots,x_n)} = x_{\lambda(n,k;x_1,\ldots,x_n)}$$

$$= \prod_{k=1}^{n} G_{x_{\lambda(n,k;x_1,\ldots,x_n)}}(y_k),$$

and that completes the proof.

Lemma 2.

Under condition 3, $\sup_{0\le t\le 1} |\psi_n(t) - \psi(t)| \to 0$ a.s.

Proof:

Let $V(\sigma^2)$ denote the total variation of $\sigma^2(x)$ on $(-\infty, \infty)$. By condition 3, $V(\sigma^2) < \infty$. Condition 3 also implies that $\sigma^2(x)$ is bounded on $(-\infty, \infty)$. Let M denote this bound. Since $\psi_n(t) = 0$ for $0 \le t < 1/n$,

$$\sup_{0\le t<1/n} |\psi_n(t) - \psi(t)| = \sup_{0\le t<1/n} \psi(t) \le \psi(1/n),$$

and $\lim_{n\to\infty} \psi(1/n) = 0$. It therefore suffices to show that $\sup_{1/n\le t\le 1} |\psi_n(t) - \psi(t)| \to 0$ a.s. Now

$$\sup_{1/n\le t\le 1} |\psi_n(t) - \psi(t)| = \sup_{1/n\le t\le 1} \left|\int_{-\infty}^{X_{n,[nt]}} \sigma^2(x)dF_n(x) - \int_{-\infty}^{F^{-1}(t)} \sigma^2(x)dF(x)\right|$$

$$\le \sup_{1/n\le t\le 1} \left|\int_{-\infty}^{X_{n,[nt]}} \sigma^2(x)d[F_n(x)-F(x)]\right|$$

$$+ \sup_{-\infty<x<\infty} \sigma^2(x) \sup_{1/n\le t\le 1} |F(X_{n,[nt]}) - t|$$

$$\le \sup_{-\infty<u<\infty} \left|\int_{-\infty}^{u} \sigma^2(x)d[F_n(x)-F(x)]\right| + M \sup_{1/n\le t\le 1} |F(X_{n,[nt]}) - t|.$$

Hence it suffices to show that

(5) $$\sup_{-\infty<u<\infty} \left|\int_{-\infty}^{u} \sigma^2(x)d[F_n(x)-F(x)]\right| \to 0 \quad \text{a.s.}$$

and

(6) $$\sup_{1/n\le t\le 1} |F(X_{n,[nt]}) - t| \to 0 \quad \text{a.s.}$$

as $n \to \infty$. Integrating by parts, we have

$$\sup_{-\infty<u<\infty} \left|\int_{-\infty}^{u} \sigma^2(x)d[F_n(x)-F(x)]\right|$$

$$= \sup_{-\infty<u<\infty} \left|-\int_{-\infty}^{u} [F_n(x)-F(x)]d\sigma^2(x) + \sigma^2(u)[F_n(u)-F(u)]\right|$$

$$\leq \sup_{-\infty<u<\infty} \int_{-\infty}^{u} |F_n(x)-F(x)|\cdot|d\sigma^2(x)|$$

$$+ \sup_{-\infty<u<\infty} \sigma^2(u) \sup_{-\infty<u<\infty} |F_n(u)-F(u)|$$

$$\leq \{V(\sigma^2) + M\} \sup_{-\infty<u<\infty} |F_n(u)-F(u)|,$$

and (5) follows from the Glivenko-Cantelli theorem. Finally,

$$\sup_{1/n\leq t\leq 1} |F(X_{n,[nt]}) - t| \leq \sup_{1/n\leq t\leq 1} |F_n(X_{n,[nt]}) - F(X_{n,[nt]})|$$

$$+ \sup_{1/n\leq t\leq 1} |F_n(X_{n,[nt]}) - t|$$

$$\leq \sup_{-\infty<x<\infty} |F_n(x)-F(x)| + \sup_{1/n\leq t\leq 1} |[nt]/n - t|$$

$$\leq \sup_{-\infty<x<\infty} |F_n(x)-F(x)| + \frac{1}{n} ,$$

and another application of the Glivenko-Cantelli theorem leads to (6).

We now proceed to study the asymptotic behavior of

$$\{S_{n,[nt]}/\sqrt{n\psi(1)} , 0 \leq t \leq 1\}$$

where $S_{nk} = \sum_{j=1}^{k} \{Y_{nj} - m(X_{nj})\}$, $k = 1,\ldots, n$ and $\psi(t)$ is given by (1). For this we need a bound for the second moment of the first exit time of a Brownian Motion. Such bounds have been obtained for moments of all orders by Skorokhod ((1961), p. 166). We give an alternative derivation of the bound for the second moment which may be of some interest in itself.

3. First Exit of a Brownian Motion.

Consider the stopping time T when a Brownian Motion $\{\xi(t), t \geq 0\}$ first escapes from the interval (a, b) for some $a < 0 < b$. It is well-known (see, e.g., Proposition 13.5 of Breiman (1968) that $E\,\xi(T) = 0$ and $E\,\xi^2(T) = ET = |ab|$. We shall now derive a formula for ET^2 in the following lemma. The formula for ET is also stated in this theorem for easy reference.

Lemma 3.

Let $\{\xi(t), t \geq 0\}$ be a separable Brownian Motion, $a < 0 < b$, and

$$T = \inf\{t \mid \xi(t) \notin (a, b)\}.$$

Then

(a) $ET = E\xi^2(T) = |ab|$

(b) $ET^2 = 2\{E[T\xi^2(T)] - E\int_0^T \xi^2(t)dt\}.$

Proof:

Besides sample path continuity we use another important property of a separable Brownian Motion, viz. joint measurability of $\xi(t, \omega)$. As a consequence of this latter property, $E\int_\alpha^\beta \xi(t)dt$ and $E\int_\alpha^\beta \xi^2(t)dt$ exist for all $0 \leq \alpha < \beta$ and can be evaluated by interchanging the order of integration. For these facts, the reader may consult Doob (1953), Chapter II, Theorems 2.5 and 2.7. Our proof employs the standard method of truncation and discretization of the stopping time T. Fix $\tau > 0$, a binary rational. Let $T^* = \min(T, \tau)$, and

$$T_N = \min\{i2^{-N} \mid i2^{-N} \geq T^*\}.$$

Then for sufficiently large N, T_N takes values in a countable set $\{t_j\}$ in $[0, \tau]$. (For small N, T_N may be as large as $\tau + 2^{-N}$). For each N, T_N is a stopping time and $T_N \downarrow T^*$ as $N \to \infty$. We shall prove (b), first for T_N, then for T^* by allowing $N \to \infty$, and finally for T by allowing $\tau \to \infty$.

Fix N. Let $B_j = \{T_N = t_j\}$. Then

$$E\int_{T_N}^{\tau} \xi^2(t)dt = \sum_j E[I_{B_j}\int_{t_j}^{\tau} \xi^2(t)dt] = \sum_j E[I_{B_j}\int_{t_j}^{\tau}\{\xi(t_j) + (\xi(t)-\xi(t_j))\}^2 dt]$$

$$= \sum_j E[I_{B_j}\int_{t_j}^{\tau} \xi^2(t_j)dt] + \sum_j E[I_{B_j}\int_{t_j}^{\tau} (\xi(t)-\xi(t_j))^2 dt]$$

$$+ 2\sum_j E[I_{B_j}\int_{t_j}^{\tau} \xi(t_j)(\xi(t)-\xi(t_j))dt] .$$

We now examine the three terms of the last expression separately, using the joint measurability and strong Markov property of the Brownian Motion.

$$\sum_j E[I_{B_j}\int_{t_j}^{\tau} \xi^2(t_j)dt] = \sum_j E[I_{B_j}\ \xi^2(t_j)(\tau-t_j)] = \tau\, E\xi^2(T_N) - E[T_N\xi^2(T_N)],$$

$$\sum_j E[I_{B_j}\int_{t_j}^{\tau} (\xi(t)-\xi(t_j))^2 dt] = \sum_j EI_{B_j}\ E\int_{t_j}^{\tau}(\xi(t) - \xi(t_j))^2 dt$$

$$= \sum_j EI_{B_j}\int_{t_j}^{\tau} E[(\xi(t)-\xi(t_j))^2]dt = \sum_j EI_{B_j}\int_{t_j}^{\tau}(t-t_j)dt$$

$$= \sum_j EI_{B_j}\ \tfrac{1}{2}(\tau-t_j)^2 = \tfrac{1}{2}\, E(\tau-T_N)^2,$$

and for each j,

$$E[I_{B_j}\int_{t_j}^{\tau} \xi(t_j)(\xi(t)-\xi(t_j))dt] = E[I_{B_j}\ \xi(t_j)]E\int_{t_j}^{\tau} (\xi(t)-\xi(t_j))dt$$

$$= E[I_{B_j}\ \xi(t_j)]\int_{t_j}^{\tau} E(\xi(t)-\xi(t_j))dt = 0.$$

Thus

$$E\int_{T_N}^{\tau} \xi^2(t)dt = \tau\, E\xi^2(T_N) - E[T_N\xi^2(T_N)] + \tfrac{1}{2}E(\tau-T_N)^2$$

$$\therefore\ \tfrac{1}{2}\tau^2 = E\int_0^{\tau}\xi^2(t)dt = E\int_0^{T_N}\xi^2(t)dt + \tau\, E\xi^2(T_N) - E[T_N\xi^2(T_N)] + \tfrac{1}{2}E(\tau-T_N)^2$$

$$= E\int_0^{T_N}\xi^2(t)dt + \tau[E\xi^2(T_N) - ET_N] + \tfrac{1}{2}\tau^2 + \tfrac{1}{2}ET_N^2 - E[T_N\xi^2(T_N)].$$

Since $E\xi^2(T_N) = ET_N$ (this holds for any stopping time taking values in a countable subset of a finite interval), we have

$$ET_N^2 = 2\{E[T_N\xi^2(T_N)] - E\int_0^{T_N}\xi^2(t)dt\},$$

and (b) is proved for T_N. We now proceed to the limit as $N\to\infty$. We have already noted that $T_N \downarrow T^* = \min(T,\tau)$. Furthermore, by path continuity, $\xi^2(T_N)\to\xi^2(T^*)$ and $\int_0^{T_N}\xi^2(t)dt \to \int_0^{T^*}\xi^2(t)dt$ with probability 1. Also,

$$T_N^2 \le \tau^2,$$

$$T_N\xi^2(T_N) \le \tau[\sup_{t\le\tau}|\xi(t)|]^2 = \tau[\max\{\sup_{t\le\tau}\xi(t),\ \inf_{t\le\tau}\xi(t)\}]^2$$

$$\le \tau\max\{[\sup_{t\le\tau}\xi(t)]^2,\ [\inf_{t\le\tau}\xi(t)]^2\},$$

and $\int_0^{T_N}\xi^2(t)dt \le \int_0^{\tau}\xi^2(t)dt$. Since

$$E[\sup_{t\le\tau}\xi(t)]^2 = E[\inf_{t\le\tau}\xi(t)]^2 = \frac{2}{\sqrt{2\pi\tau}}\int_0^{\infty}x^2e^{-x^2/2\tau}dx = \tau < \infty$$

and $E\int_0^{\tau}\xi^2(t)dt = \frac{1}{2}\tau^2 < \infty$, (b) holds for T^* by the dominated convergence theorem. We now let $\tau\to\infty$. If $T\le\tau$,

$$T^*\xi^2(T^*) = T\xi^2(T) = \max(a^2, b^2)T$$

and

$$\int_0^{T^*} \xi^2(t)dt = \int_0^{T} \xi^2(t)dt \le \int_0^{T} \max(a^2, b^2)dt = \max(a^2, b^2)T$$

if $T > \tau$,

$$T^*\xi^2(T^*) = \tau\xi^2(\tau) \le \max(a^2, b^2)\tau \le \max(a^2, b^2)T$$

and

$$\int_0^{T^*} \xi^2(t)dt = \int_0^{\tau} \xi^2(t)dt \le \int_0^{\tau} \max(a^2, b^2)dt = \max(a^2, b^2)\tau \le \max(a^2, b^2)T.$$

Thus, both $T^*\xi^2(T^*)$ and $\int_0^{T^*} \xi^2(t)dt$ are bounded by $\max(a^2, b^2)T$ and we know that $ET < \infty$. Hence by the dominated convergence theorem,

$$E[T^*\xi^2(T^*)] \to E[T\xi^2(T)]$$

and

$$E\int_0^{T^*} \xi^2(t)dt \to E\int_0^{T} \xi^2(t)dt$$

as $\tau \to \infty$. Furthermore, as $\tau \to \infty$, $T^{*2} \uparrow T^2$, so

$$ET^{*2} \to ET^2$$

by the monotone convergence theorem, and that concluds the proof.

Skorokhod's bound for the second moment of T follows as an immediate corollary from the above lemma.

Corollary.

In the framework of Lemma 3,

$$ET^2 \le 2|ab| \max(a^2, b^2).$$

4. A Skorokhod-type Embedding of $\{S_{nk}\}$ on Brownian Motion Paths.

Skorokhod (1961) developed a method of representing the cumulative sums of independent random variables by a Brownian Motion evaluated at random times. In this section we extend this idea to represent $\{S_{nk}, k=1,\dots, n\}$.

Let $(\Omega_1, \mathcal{F}_1, P_1)$ denote the probability space of $((X_1, Y_1), (X_2, Y_2), \ldots)$, $(\Omega_2, \mathcal{F}_2, P_2)$ the probability space of a Brownian Motion $\{\xi(t), t \geq 0\}$ in its sample path formulation, and consider the probability space $(\Omega, \mathcal{F}, P) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, P_1 \times P_2)$. Then in an obvious manner $(X_1, Y_1), (X_2, Y_2), \ldots$ and $\{\xi(t), t \geq 0\}$ can be thought of as random variables on $(\Omega, \mathcal{F}, P)$. In this framework, $(X_1, Y_1), (X_2, Y_2), \ldots$ are still independent random vectors following their original common distribution and $\{\xi(t), t \geq 0\}$ is still a Brownian Motion independent of (X_1, Y_1), $(X_2, Y_2), \ldots$. Let $\mathcal{G} \subset \mathcal{F}$ denote the σ-field of subsets of Ω induced by $\{X_n, n = 1,2,\ldots\}$. From now on we work in this set-up. For two stochastic processes we write $\{X(t)\} \overset{d}{=} \{Y(t)\}$ to indicate the processes have the same distribution.

Theorem 1.

If condition 1 holds and if $\beta(x) = E[\{Y-m(x)\}^4 | X = x]$ exist for all x, then for every n, there exist stopping times $T_{n1}, \ldots, T_{nn}$ of the Brownian Motion $\{\xi(t), t \geq 0\}$ such that

(a) $(S_{n1}, \ldots, S_{nn}) \overset{d}{=} (\xi(T_1), \ldots, \xi(T_1 + \ldots + T_n))$.

(b) $T_{n1}, \ldots, T_{nn}$ are conditionally independent given $\mathcal{G}$.

(c) $E[T_{nk} | \mathcal{G}] = \sigma^2(X_{nk})$.

(d) $E[T_{nk}^2 | \mathcal{G}] \leq C\beta(X_{nk})$, where C is a constant.

Proof:

Argue conditionally given $\mathcal{G}$ in the probability space $(\Omega, \mathcal{F}, P)$. Clearly, $\{\xi(t), t \geq 0\}$ is still a Brownian Motion and is independent of $Y_{nk} - m(X_{nk})$, $k = 1,\ldots, n$. Furthermore, by Lemma 1, the random variables $Y_{nk} - m(X_{nk})$, $k = 1,\ldots, n$ are mutually independent with mean 0, variance $\sigma^2(X_{nk})$ and fourth moment $\beta(X_{nk})$. The theorem is now proved by repeated applications of Skorokhod's (1961) Lemma 3 of Chapter 7, his Remarks 1 and 2 following this lemma (pp. 167-168), and our Lemma 3.

5. Convergence of Sample Paths of $\{S_{nk}\}$ and an Invariance Principle.

By means of the embedding theorem of the last section we now study the convergence of normalized cumulative sums of induced order statistics. The following is the main theorem of this section.

Theorem 2.

Under conditions 1-3, there exist processes $\{\xi^{(n)}(t),\ 0 \le t \le 1\}$ and a Brownian Motion $\{\xi(t),\ t \ge 0\}$ on a common probability space such that

(a) for each n,

$$\{\xi^{(n)}(t),\ 0 \le t \le 1\} \overset{d}{=} \{S_{n,[nt]}/\sqrt{n\psi(1)}\ ,\ 0 \le t \le 1\},$$

(b) for any sufficiently rapidly increasing subsequence $\{n_j\}$

$$\sup_{0\le t\le 1} |\xi^{(n_j)}(t) - \xi(\psi(t)/\psi(1))| \to 0 \quad \text{a.s.}$$

Proof:

We shall prove the theorem in the context of the probability space $(\Omega, \mathfrak{F}, P)$ and the Brownian Motion $\{\xi(t),\ t \ge 0\}$ described at the beginning of Section 4. For each n, construct random stopping times $T_{n1}, \ldots, T_{nn}$ of $\xi(t)$ as in Theorem 1. Then for each n,

$$\{S_{n,[nt]}/\sqrt{n\psi(1)}\ ,\ 0 \le t \le 1\} \overset{d}{=} \{\frac{1}{\sqrt{n\psi(1)}}\ \xi(T_{n1}+\ldots+T_{n,[nt]}), 0\le t\le 1\}$$

$$\overset{d}{=} \{\xi(\frac{T_{n1}+\ldots+T_{n,[nt]}}{n\psi(1)}),\ 0\le t\le 1\}.$$

Thus the processes

$$\{\xi^{(n)}(t) = \xi(\frac{T_{n1}+\ldots+T_{n,[nt]}}{n\psi(1)}),\ 0 \le t \le 1\}$$

satisfy (a). We shall now show that these processes also satisfy (b). Arguing as in the proof of Theorem 13.8 of Breiman (1968), it will suffice to show that

$$\sup_{0\le t\le 1} \left| n^{-1} \sum_{k=1}^{[nt]} T_{nk} - \psi(t) \right| \xrightarrow{p} 0$$

as $n \to \infty$. Now

$$\sup_{0\le t\le 1} \left| n^{-1} \sum_{k=1}^{[nt]} T_{nk} - \psi(t) \right| \le \sup_{0\le t\le 1} \left| n^{-1} \sum_{k=1}^{[nt]} T_{nk} - \psi_n(t) \right|$$

$$+ \sup_{0\le t\le 1} |\psi_n(t) - \psi(t)|,$$

and the second term in the last expression converges to 0 a.s. by Lemma 2. It therefore remains to be shown that the first term converges to 0 in probability. Now for any $\epsilon > 0$ and $n > 1/\epsilon$,

(7) $$\sup_{0\le t\le 1} \left| n^{-1} \sum_{k=1}^{[nt]} T_{nk} - \psi_n(t) \right| = \sup_{1/n\le t\le 1} \left| n^{-1} \sum_{k=1}^{[nt]} \{T_{nk} - \sigma^2(X_{nk})\} \right|$$

$$= \sup_{1/n\le t\le 1} (t/[nt]) \left| \sum_{k=1}^{[nt]} \{T_{nk} - \sigma^2(X_{nk})\} \right|$$

$$\le \epsilon \sup_{1\le k\le [\epsilon n]} \left| k^{-1} \sum_{j=1}^{k} \{T_{nj} - \sigma^2(X_{nj})\} \right| + \sup_{[\epsilon n]\le k\le n} \left| k^{-1} \sum_{j=1}^{k} \{T_{nj} - \sigma^2(X_{nj})\} \right|$$

We now apply Theorem 1(b), (c), (d), the Hájek-Rényi (1955) inequality, and use condition 2 to get

$$P[\sup_{1\le k\le [\epsilon n]} \left| k^{-1} \sum_{j=1}^{k} \{T_{nj} - \sigma^2(X_{nj})\} \right| > x | G] \le CBx^{-2} \sum_{k=1}^{[\epsilon n]} k^{-2}$$

and therefore,

(8) $$P[\sup_{1<k\le [\epsilon n]} \left| k^{-1} \sum_{j=1}^{k} \{T_{nj} - \sigma^2(X_{nj})\} \right| > x] \le CBx^{-2} \sum_{k=1}^{[\epsilon n]} k^{-2},$$

and

$$P[\sup_{[\epsilon n]\le k\le n} \left| k^{-1} \sum_{j=1}^{k} \{T_{nj} - \sigma^2(X_{nj})\} \right| > x | G] \le CBx^{-2} n([\epsilon n])^{-2},$$

and therefore,

$$(9) \qquad P[\sup_{[\varepsilon n]\le k\le n} |k^{-1}\sum_{j=1}^{k}\{T_{nj}-\sigma^2(X_{nj})\}| > x] \le CBx^{-2}n\{[\varepsilon n]\}^{-2}.$$

From (7), (8) and (9) we have for any $\delta > 0$ and $\varepsilon > 0$,

$$\overline{\lim_{n\to\infty}} P[\sup_{0\le t\le 1} |n^{-1}\sum_{k=1}^{[nt]} T_{nk} - \psi_n(t)| > \delta]$$

$$\le \overline{\lim_{n\to\infty}} P[\varepsilon \sup_{1\le k\le [\varepsilon n]} |k^{-1}\sum_{j=1}^{k}\{T_{nj}-\sigma^2(X_{nj})\}| > \delta/2]$$

$$+ \overline{\lim_{n\to\infty}} P[\sup_{[\varepsilon n]\le k\le n} |k^{-1}\sum_{j=1}^{k}\{T_{nj}-\sigma^2(X_{nj})\}| > \delta/2]$$

$$= \overline{\lim_{n\to\infty}} P[\sup_{1\le k\le [\varepsilon n]} |k^{-1}\sum_{j=1}^{k}\{T_{nj}-\sigma^2(X_{nj})\}| > \delta/2\varepsilon] + 0$$

$$\le 4CB\varepsilon^2\delta^{-2}\sum_{k=1}^{\infty} k^{-2},$$

which goes to 0 for any given $\delta > 0$ by allowing ε to tend to zero. This concludes the proof.

An immediate consequence of Theorem 2 is the following invariance principle. Let D denote the space of all functions on $[0, 1]$ which are continuous at 0 and 1 and are right-continuous and have left limits on $(0, 1)$.

Theorem 3.

Let H defined on D be continuous in the topology of uniform convergence. Let $\{\zeta^{(n)}(t) = S_{n,[nt]}/\sqrt{n\psi(1)}\ ,\ 0 \le t \le 1\}$ and $\{\zeta(t) = \xi(\psi(t)/\psi(1)),\ 0 \le t \le 1\}$ where $\{\xi(t),\ t \ge 0\}$ is a Brownian Motion. Then

$$H(\zeta^{(n)}(\cdot)) \to H(\zeta(\cdot))$$

in distribution as $n \to \infty$.

In particular, for $H_1(x(\cdot)) = \sup_{0\le t\le 1} x(t)$ and $H_2(x(\cdot)) = \sup_{0\le t\le 1} |x(t)|$ on D, we have the following corollary.

Corollary.

(a) $$\lim_{n\to\infty} P[\sup_{0\le t\le 1} S_{n,[nt]}/\sqrt{n\psi(1)} \le \lambda] = \int_{-\lambda}^{\lambda} (2\pi)^{-\frac{1}{2}} e^{-x^2/2} dx$$

(b) $$\lim_{n\to\infty} P[\sup_{0\le t\le 1} |S_{n,[nt]}|/\sqrt{n\psi(1)} \le \lambda] = \sum_{k=-\infty}^{\infty} (-1)^k \int_{(2k-1)\lambda}^{(2k+1)\lambda} (2\pi)^{-\frac{1}{2}} e^{-x^2/2} dx.$$

Proof:

By Theorem 3, $\sup_{0\le t\le 1} S_{n,[nt]}/\sqrt{n\psi(1)}$ has the same limiting distribution as $\sup_{0\le t\le 1} \xi(\psi(t)/\psi(1))$. Since F is continuous, $\psi(t)/\psi(1)$ increases continuously from 0 to 1 as t increases from 0 to 1. Thus

$$\{\psi(t)/\psi(1) \mid 0 \le t \le 1\} = [0, 1].$$

Hence

$$\sup_{0\le t\le 1} \xi(\psi(t)/\psi(1)) = \sup_{0\le t\le 1} \xi(t),$$

and (a) follows from well-known results. The proof of (b) is exactly similar. (See, e.g., Freedman (1971), Corollary 29 and Proposition 34.)

For simplicity, we have stated Theorem 3 for continuous H. However, this invariance principle holds for a more general class of functions (see Breiman (1968), Theorem 13.12).

6. Testing a Specified Regression Function.

Using the results of the last section, we can construct tests for a specified regression function. We want to test the null hypothesis that the regression function $m(x)$ of Y on X in a bivariate distribution is equal to a specified function $m_0(x)$. Let $(X_1, Y_1),\ldots, (X_n, Y_n)$ be independent samples from this distribution. We then compute the order statistics $X_{n1},\ldots, X_{nn}$ of the X-observations and the induced order statistics $Y_{n1},\ldots, Y_{nn}$ of the Y observations, and let

$$S_{nk} = \sum_{j=1}^{k} \{Y_{nj} - m_0(X_{nj})\}.$$

Then under the null hypothesis,

$$(10) \qquad P[\max_{k=1,\ldots,n} |S_{nk}|/\sqrt{n\psi(1)} \leq \lambda] = \sum_{k=-\infty}^{\infty} (-1)^k \int_{(2k-1)\lambda}^{(2k+1)\lambda} (2\pi)^{-\frac{1}{2}} e^{-x^2/2} dx.$$

However, $\psi(1) = \int_{-\infty}^{\infty} \sigma^2(x)dF(x)$ is unknown, but

$$\hat{\psi}_n(1) = n^{-1} \sum_{j=1}^{n} \{Y_{nj} - m_0(X_{nj})\}^2 = n^{-1} \sum_{j=1}^{n} \{Y_j - m_0(X_j)\}^2$$

is a consistent estimator of $\psi(1)$ and (10) holds with $\psi(1)$ replaced by $\hat{\psi}_n(1)$. We can now use the large sample level α test:

<u>Test 1</u>.

Reject the null hypothesis if and only if

$$\max_{k=1,\ldots,n} |S_{nk}|/\sqrt{n\hat{\psi}_n(1)} \geq \lambda_\alpha$$

where $\sum_{k=-\infty}^{\infty} (-1)^k \int_{(2k-1)\lambda_\alpha}^{(2k+1)\lambda_\alpha} (2\pi)^{-\frac{1}{2}} e^{-x^2/2} dx = 1 - \alpha.$

The function $H(x(\cdot)) = \int_0^1 x(t)dt$ is also continuous on D and the invariance principle applies to the asymptotic distribution of

$$\{n\psi(1)\}^{-\frac{1}{2}} \int_0^1 S_{n,[nt]}\, dt.$$

Thus under the null hypothesis, $\{n\psi(1)\}^{-\frac{1}{2}} \int_0^1 S_{n,[nt]}dt$ converges in distribution to $\int_0^1 \xi(\psi(t)/\psi(1))dt$ where $\xi(t)$ is a Brownian Motion. It is easily seen that $\int_0^1 \xi(\psi(t)/\psi(1))dt$ is a normal random variable with mean 0 and variance $\{\psi(1)\}^{-1} \int_0^1\int_0^1 \psi(\min(s, t))ds\, dt$. Hence under the null hypothesis,

$$\int_0^1 S_{n,[nt]}dt/[n\int_0^1\int_0^1 \psi(\min(s, t))dsdt]^{\frac{1}{2}}$$

is asymptotically normally distributed with mean 0 and variance 1. The function $\psi(t)$ can be estimated from the sample by

$$\hat{\psi}_n(t) = n^{-1}\sum_{k=1}^{[nt]}\{Y_{nk} - m_0(X_{nk})\}^2.$$

To see that $\hat{\psi}_n(t)$ is a uniformly consistent estimate of $\psi(t)$, note that

$$\sup_{0\le t\le 1}|\hat{\psi}_n(t) - \psi(t)| \le \sup_{0\le t\le 1}|\hat{\psi}_n(t) - \psi_n(t)| + \sup_{0\le t\le 1}|\psi_n(t) - \psi(t)|$$

where $\psi_n(t)$ is as defined in (2). By Lemma 2, $\sup_{0\le t\le 1}|\psi_n(t) - \psi(t)| \to 0$ a.s., and it can be shown in a way analogous to the proof of Theorem 3, that $\sup_{0\le t\le 1}|\hat{\psi}_n(t) - \psi_n(t)| \xrightarrow{P} 0$. Hence,

$$\int_0^1\int_0^1 \hat{\psi}_n(\min(s, t))dsdt \xrightarrow{P} \int_0^1\int_0^1 \psi(\min(s, t)dsdt,$$

and consequently,

$$W_n = \int_0^1 S_{n,[nt]}dt \;/\; [n\int_0^1\int_0^1 \hat{\psi}_n(\min(s, t))dsdt]^{\frac{1}{2}}$$

is also asymptotically normally distributed with mean 0 and variance 1 under the null hypothesis. We can now use the following large sample level α tests:

Test 2a.

Reject the null hypothesis if and only if

$$W_n \ge \Phi^{-1}(1-\alpha),$$

or

Test 2b.

Reject the null hypothesis if and only if

$$W_n \le \Phi^{-1}(\alpha),$$

where Φ is the cdf of a normal random variable with mean 0 and variance 1.

By a little algebraic simplification, we have

$$W_n = \sum_{j=1}^{n-1} (n-R_{nj})\{Y_j - m_0(X_j)\} \Big/ \Big[\sum_{j=1}^{n-1} (n^2 - R_{nj}^2)\{Y_j - m_0(X_j)\}^2\Big]^{\frac{1}{2}},$$

where R_{nj} is the rank of X_j among $X_1, \ldots, X_n$. In this form, W_n is computed easily.

Test 1 would guard against all possible alternatives, whereas Tests 2a and 2b would guard against alternatives $m(x) > m_0(x)$ and $m(x) < m_0(x)$ respectively.

REFERENCES

[1] Breiman, Leo (1968). <u>Probability</u>. Addison-Wesley, Reading, Massachusetts.

[2] Donsker, M. (1951). An invariance principle for certain probability limit theorems. <u>Mem</u>. <u>Amer</u>. <u>Math</u>. <u>Soc</u>., No. 6.

[3] Doob, J. L. (1953). <u>Stochastic processes</u>. John Wiley & Sons, New York.

[4] Freedman, David (1971). <u>Brownian motion and diffusion</u>. Holden-Day, San Francisco.

[5] Hájeck, J. and Rényi, A. (1955). Generalization of an inequality of Kolmogorov. <u>Acta</u> <u>Math</u>. <u>Acad</u>. <u>Sci</u>. <u>Hungar</u>. <u>6</u> 281-283.

[6] Skorokhod, A. V. (1961). <u>Studies in the theory of random processes</u>. Kiev University Press. (English Translation (1965). Addison-Wesley, Reading, Massachusetts. Page numbers referred to in the text are from the English Translation.)